

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Wavelet Based Speech Strategy in Cochlear Implant

Gulden Kokturk
Dokuz Eylul University
Turkey

1. Introduction

A significant percentage of the populations in developed countries encounter hearing impairment. Cochlear Implant was developed to increase the hearing capacity of these people. In recent years, adults and children have benefited from the usage of Cochlear Implant and the users were positively affected by the improvement of implant techniques. Although these devices allow for increased performance, a significant gap still remains in speech recognition between Cochlear Implant user and people, who possess normal listening capabilities.

The Cochlear implant prosthesis, works with direct stimulation of the auditory nerve cells of deaf people, whose recipient cells in their cochlea were destroyed (House&Berliner, 1982; House&Urban, 1979; Christiansen&Leigh, 2002; Loeb, 1990; Parkins&Anderson, 1983; Wilson, 1993). The system basically consists of the following sections; microphone, speech processor, transmitter, receiver and electrode array. Mainly, implant stimulation systems, signal processing strategies and techniques in increasing the suitability of the instrument to various patients were the improved upon this prosthesis. However, new developments in this area also exist on the configuration of hearing function, especially signal processing strategies.

Speech processing techniques are very important in increasing the users' hearing potential (Eddington, 1980; Wilson et al., 1991; Dormon et al., 1997; Moore&Teagle, 2002). A variety of strategies were developed in the recent years with the aim to improve the hearing abilities of deaf people and take these abilities closer to those of people with natural hearing. Of these, various speech processing strategies were developed for multi-channel cochlear implants (Derbel et al., 1994; Cheikhrouhou et al., 2004; Gopalakrishna et al., 2010; Millar et al., 1984). These strategies can be classified mainly in three parts; waveform strategies, feature extraction strategies and hybrid strategy (N-of-M strategy).

Wavelet method is a basic method that is used for noise filtering, compression and analysis of nonstationary signals and images. The wavelet transform is an appropriate method for semistationary signals and provides a good resolution in both time and frequency. Several studies were carried out on the use of the wavelet transform for speech processing. The wavelet transform gives better results than traditional methods in improving speech. The wavelet packet transform is a type of the discrete wavelet transform that allows for subband analysis in the second decomposition without any constraints. Basically, the wavelet packet

transform adapts to the frequency axis and allows decomposition. This particular decomposition is implemented with the optimization criteria. Studies show that the filter banks used in the sound strategies of cochlear implants are parallel to the filter banks used in the wavelet transform. This indicates that the wavelet transform is a method that can be used in Cochlear implants. However, the author could not find any distinctive study on the analysis of speech using the wavelet packet transform for speech processing.

In this study, a new strategy, which uses the wavelet packet transform in obtaining the sequence of the electrode array to be stimulated, is proposed. Thus the stimulated electrode array is based upon but would be different from the N-of-M Strategy. An experimental study was performed and analyzed using Turkish words. The words were selected and grouped based on Turkish sound knowledge, and were recorded in the silent rooms of Ege University, Department of Otology. The experimental subjects listened to the words for both the N-of-M strategy and the proposed strategy, thus comparative results were obtained. Better intelligibility results were attained by the usage of the proposed strategy.

2. Strategies for cochlear implants

Multi-channel implants provide electrical stimulation in the cochlea using an array of electrodes. An electrode array is used so that different auditory nerve fibers can be stimulated at different places in the cochlea. Electrodes respond for each frequency of the signal and hair cells. The ones near the base of the cochlea are stimulated with high frequency signals while electrodes near the apex are stimulated with low frequency signals.

The various signal processing strategies developed for multi-channel cochlea can be examined under two main categories: waveform strategies and feature-extraction strategies. These strategies extract the speech information from the speech signal and redeliver it to the electrodes. The waveform strategies use some type of waveform (in analog or pulsatile form) derived by filtering the speech signal into different frequency bands. The feature extraction strategies use some type of spectral features, such as formants, derived by feature extraction algorithms.

There are various parameters that present the acoustic signal information to the electrodes in these signal processing strategies. The first parameter is the number of electrodes used for stimulus that decides the frequency resolution. Mostly, it is used as 16-22 electrodes for stimulation. This also depends on receivers of the individual cochlear implant relating to neuron population distribution. The second parameter is the electrode configuration. Different configurations are controlled since the electric current distributes electrodes symmetrically. The most important parameter is the amplitude of the electric current which is constructed using some envelope detection algorithms from the filtered waveform. It is controlled by the loudness level of the stimulation that could be comprehended. The electric current amplitude includes spectral information generated from the time varying current amplitude levels on each electrode and on different electrodes stimulated in the same cycle.

The main strategies are discussed below.

2.1 Compressed analog approach

The strategy of compressed analog approach has been developed mainly by Symbion Company (Edington, 1980). In this system, first, the audio signal is compressed using the

automatic gain control unit. Later the signal is filtered with four side by side frequency bands at the center frequencies of 0.5, 1, 2 and 3.4 kHz. Then, the filtered waveforms are changed into stimulation format and sent to four electrodes that were placed into the cochlea by surgical intervention (Dorman et al., 1989).

The compressed analog approach gives useful spectral information to the electrodes. But the channel interaction causes problems in the compressed analog approach. Since the stimulation is analog, the stimulus is transmitted continuously to the four electrodes at the same time. The simultaneous stimulation causes channel interaction and can negatively affect the performance of the device.

2.2 Continuous interleaved sampling

The continuous interleaved sampling approach was developed by researchers of the Research Triangle Institute (Wilson et al., 1991). In this strategy, the signal is sent through the electrode not by stimulation but by interleaved strokes. Here, the amplitude of the pulse is derived from the envelope of bandpass filter. Later, this resulting envelope is compressed and used for modulation of two-phase pulses. The patient’s hearing unit is electronically excited. Non-linear compression function (i.e. logarithmic function) is used to ensure that the envelope output is suitable for the patient’s dynamic range. A block diagram of the continuous interleaved strategy is shown in Figure 1.

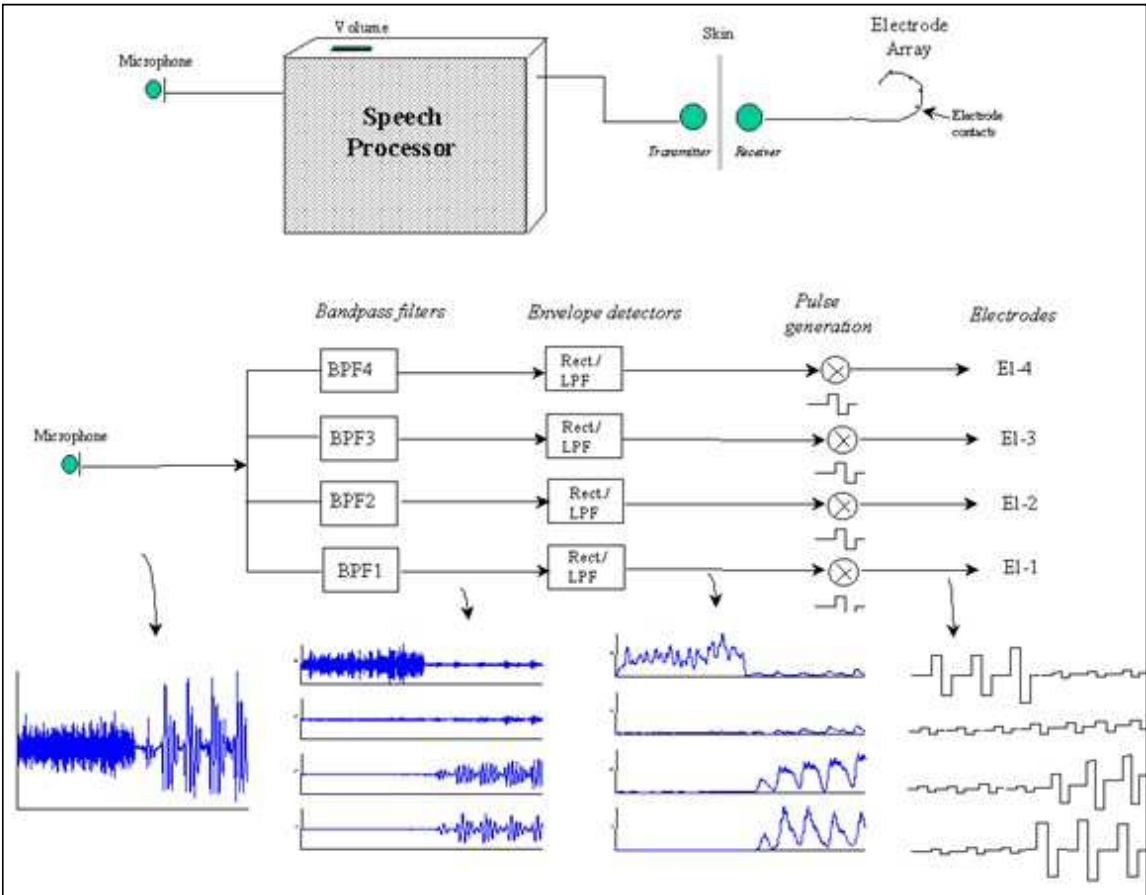


Fig. 1. Detailed block diagram of continuous interleaved sampling speech strategy in cochlear implant (Loizou, 1998)

2.3 N-of-M (N-M) speech processing

In N-M strategy, the audio signal is divided into m frequency bands; and the processor selects n number of the highest-energy envelope outputs (see in Figure 2). Only the electrodes corresponding to the selected n outputs are stimulated at each cycle (Nogueira et al., 2006). For example; in a strategy of 22-6, only 6 of the 22 channel outputs are selected, thus only these 6 selected channels are stimulated. N-M strategy is a hybrid strategy since it also includes feature representation. A general block diagram showing the feature inference for N-M system is given in Figure 3.

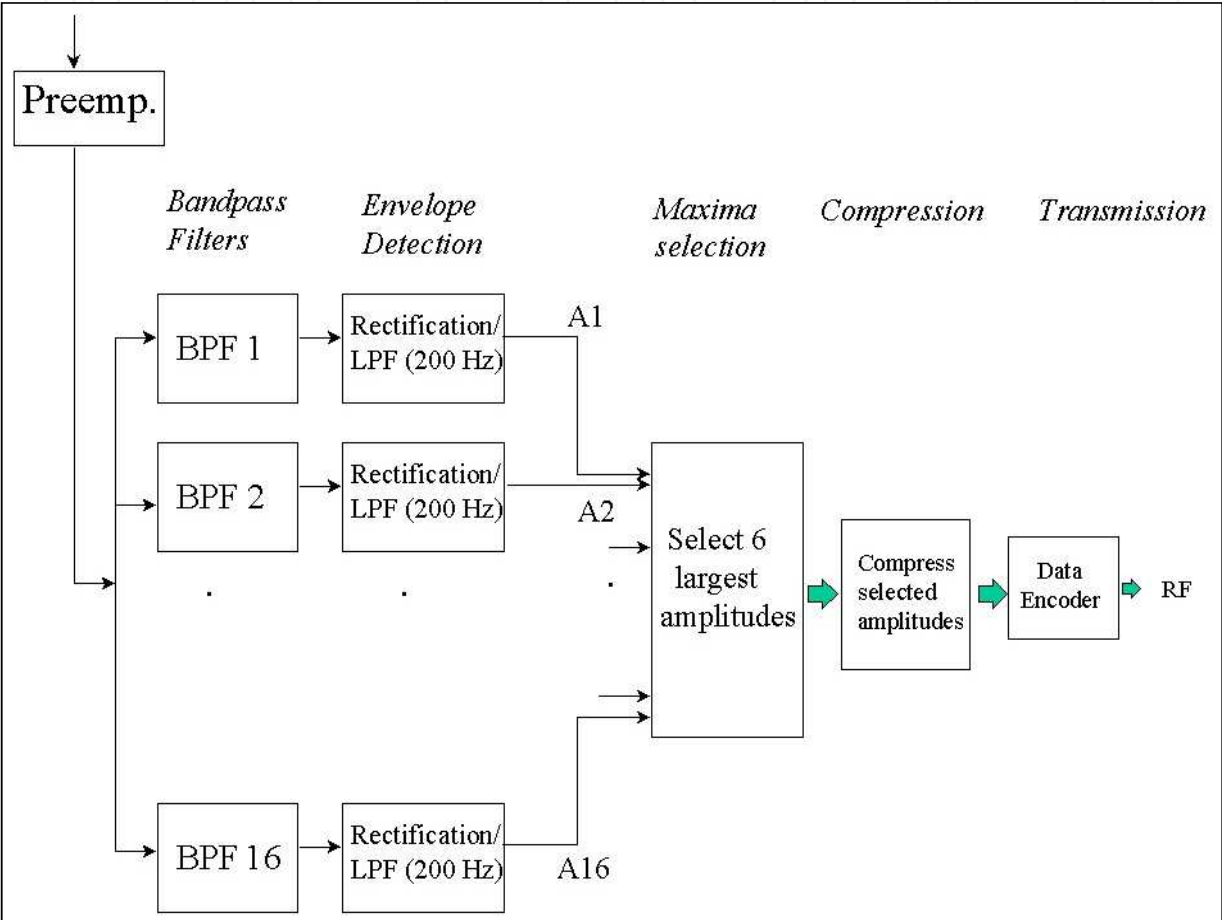


Fig. 2. Detailed block diagram of N-M speech strategy in Cochlear implant (Loizou, 1998)

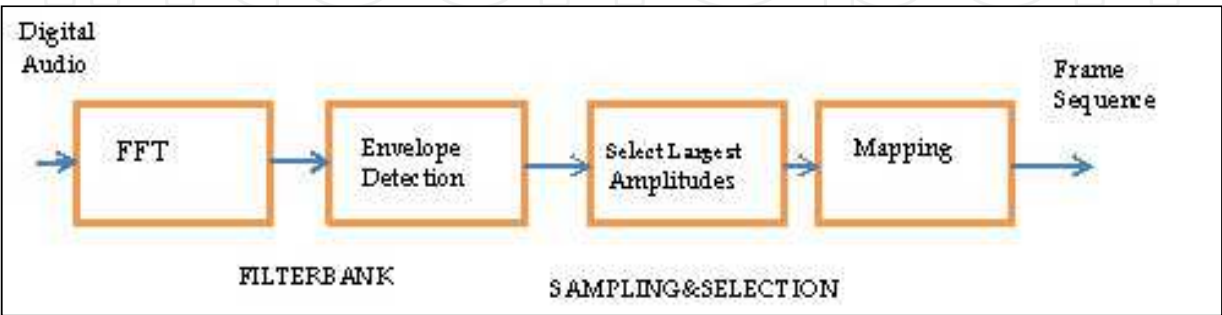


Fig. 3. General block diagram showing the feature inference of N-M speech strategy in Cochlear implant (Loizou, 1998)

3. Wavelet transform

A time signal can be evaluated by a series of coefficients, based on an analysis function. For example, a signal can be transformed from time domain to frequency domain. The oldest and best known method for this is the Fourier transform. Joseph Fourier developed his method that represents signal contents by using basis functions in 1807. Based on this work the wavelet theory was developed by Alfred Haar in 1909. In 1930's, Paul Levy improved Haar basis function using scale varying. In 1981, a transformation method of decomposing a signal into wavelet coefficients and reconstructing the original signal from these coefficients was found by Jean Morlet and Alex Grossman. And Stephane Mallat and Yves Meyer derived multiresolution decomposition using wavelets. Later, Ingrid Daubechies developed a new wavelet analysis method to construct her own family of wavelets using the multiresolution theory. The set of wavelet orthonormal basis function based on Daubechies' work is the milestone of wavelet applications today. With these developments, theoretical investigations of wavelet analysis began to accrue. (Merry, 2005).

Generally, the Fourier transform is an efficient transform for stationary and pseudo stationary signals. But this technique is not suitable for the nonstationary signals such as noisy and aperiodic signals. These signals can be analyzed using local transformation methods; the short time Fourier transform, time-frequency distributions and wavelets. All these techniques analyzed the signal using the correlation between original signal and analysis function.

Wavelet transform can be classified as; continuous wavelet transform, discrete wavelet transform and fast wavelet transform (Holschneider, 1989, Mallat, 1998; Meyer, 1992). The wavelet transform applied to a wide range of use, subjects including signal and image processing, and biomedical signal processing. The most important advantage of the wavelet transform is that it allows for the local analysis of the signal. Also, wavelet analysis reveals such as discontinuities, corruptions etc. in a signal.

A wavelet function, $\psi(t)$ is a small wave. In wavelet function, the wavelet must be zero as soon as possible while still having oscillatory. Therefore, it includes different frequencies to

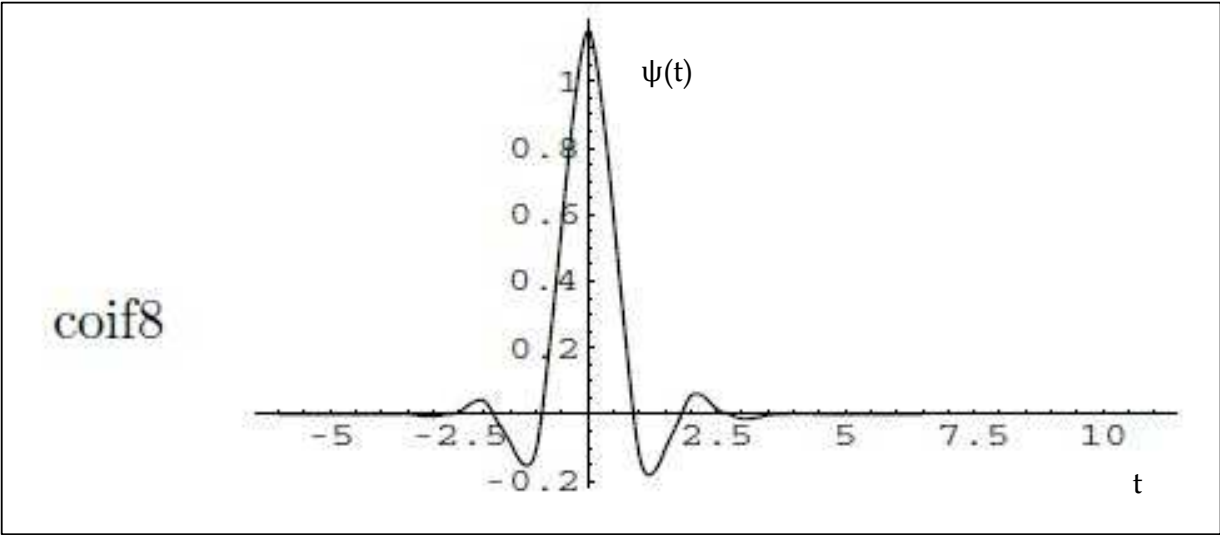


Fig. 4. An example wavelet, Coiflet wavelet

be analyzed. An example of a possible wavelet known as Coiflet wavelet function is shown in Figure 4. (Misiti et al., 2000).

The wavelet basis is a grid decomposition of the phase plane which is shown in Figure 5 (Pereyra&Mohlenkampy, 2004).

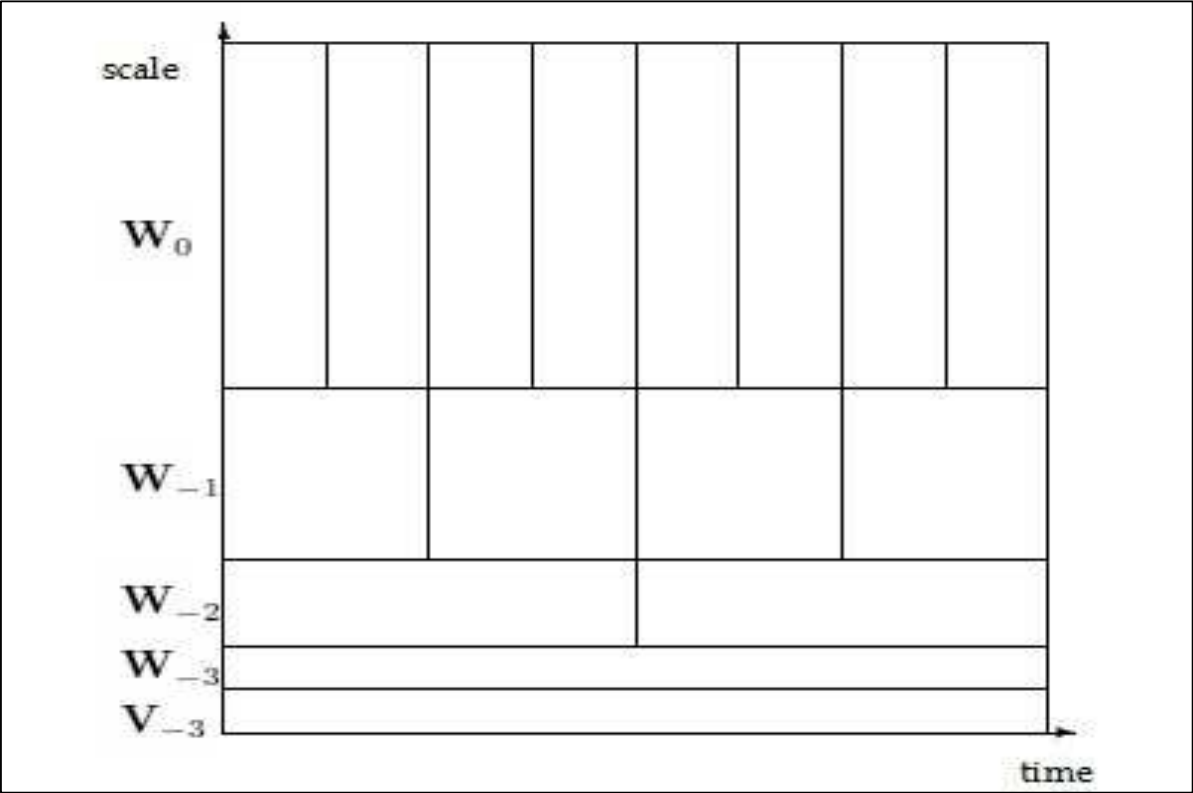


Fig. 5. The phase plane of the wavelet transform

A wavelet function $\psi(t)$ must satisfy the following conditions:

- 1. A wavelet must have finite energy

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \tag{1}$$

where E is the energy of the wavelet function. This means that the energy of the analyzing function is equal to the integrated square magnitude of $\psi(t)$ and it must have a finite value.

If $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$, then the following condition must satisfy the so called the admissibility condition.

$$C_{\psi} = \int_0^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \tag{2}$$

This condition indicates that the wavelet has no zero frequency components. The mean of wavelet must equal zero. This is known as the admissibility condition.

- 2. The Fourier transform of C_{ψ} must both be real for complex wavelets and must vanish for negative frequencies (Daubechies, 1992; Rioul&Vetterli, 1991; Pereyra & Mohlenkampy, 2004).

3.1 Continuous wavelet transform

As a family of functions, wavelets can be reconstructed from translation and dilation of a single wavelet $\psi(t)$. Here, b is the translation parameter and a is the scale parameter. It is called the mother wavelet and is defined as

$$\psi_{a,b} = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad a, b \in \mathbb{R}, \quad a \neq 0 \quad (3)$$

The integral transformation of $\psi_{a,b}$ is given by (Polikar, 1999)

$$X_{WT}(a, b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}^*(t) dt \quad (4)$$

where b is translation parameter, a is dilation parameter and $*$ indicates the complex conjugate which is used in a complex wavelet. This equation $X_{WT}(a, b)$ is called a continuous wavelet transform of $f(t)$. The energy of the signal is normalized by dividing wavelet coefficients with $1/\sqrt{|a|}$ at each scale. This enables the wavelets to have the same energy for each scale. When scale parameter is changed at each scale, the center frequency of wavelet and the window length are also changed. Therefore scale is used instead of frequency for the wavelet analysis. The translation parameter identifies the location of the wavelet function in time. Change in translation parameter shows shifted wavelet over the signal. In time-scale domain, the rows are filled for constant scale and varying translation. Similarly, the columns are filled for constant translation and varying scale. The coefficients $X_{WT}(a, b)$ are called wavelet coefficients and are associated with a scale in frequency domain and a time in time domain.

The inverse continuous wavelet transform can also be defined as

$$f(t) = \frac{1}{C_{\psi}^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_{WT}(a, b) \frac{1}{a^2} \psi\left(\frac{t-b}{a}\right) db da \quad (5)$$

Note that the admissibility constant C must adhere to the second wavelet condition.

A wavelet function has a center frequency and it changes inversely to this frequency. Small scale change indicates that high frequencies corresponding to the scale include detailed information of the signal. In contrast, a large scale corresponds to a low frequency and gives coarser information of the signal. In wavelet transform, Heisenberg inequality must be satisfied. Therefore the bandwidth of time and scale $\Delta t \Delta \omega$ is constant and bounded. When the scale is decreased the time resolution Δt will increase. This indicates that the frequency resolution $\Delta \omega$ is proportional to the frequency ω . Consequently, the wavelet transform has a constant relative frequency resolution (Debnath, 2002).

The continuous wavelet transform is a linear transformation like the Fourier transform. It accomplishes the discrete values of scale and translate parameters. The resulting coefficients are called wavelet series. Later, the discretization of coefficients can be done arbitrarily, but reconstruction is required.

3.2 Discrete wavelet transform

There are two ways to introduce the discrete wavelet. One is the discretization of the continuous wavelet transform and the other is through multiresolution analysis.

In many applications, the data is represented by a finite sequence, so it is important to analyze the discrete form of the continuous wavelet transform. From a mathematical view, a continuous notation of two continuous parameters a and b can be converted into the discrete form. Then, two positive constants a_0 and b_0 are defined as (Strang&Nguyen, 1997).

$$\psi_{m,n} = a_0^{-m/2} \psi(a_0^{-m}x - nb_0), \quad m, n \in \mathbb{Z} \quad (6)$$

It is noted that the discrete wavelet transform can be derived directly from the corresponding continuous version by using $a = a_0^m$ and $b = nb_0 a_0^m$. Therefore, the discrete wavelet transform gives a function of a finite set of the wavelet coefficients in the time-scale domain that was indexed m and n .

The aim of the multiresolution analysis is to represent a signal as a limit of consecutive approximations. These correspond to different levels of resolutions. In multiresolution analysis, the orthogonal wavelet bases are constructed using a definite set of rules, where a multiresolution analysis is a sequence of closed space V_j if the following conditions hold.

1. $\{0\} \subset \dots \subset V_1 \subset V_0 \subset V_{-1} \subset \dots \subset L^2(\mathbb{R})$
2. $f(t) \in V_j \Leftrightarrow f(2^j t) \in V_{j-1}$
3. $f(t) \in V_0 \Leftrightarrow f(t-k) \in V_0$
4. There exists a scaling function $\phi \in V_0$ such that $\{\phi(t-k)\}_{k \in \mathbb{Z}}$ is an orthonormal basis of V_0 .

The function ϕ is called the scaling function. If V_j is a multiresolution in $L^2(\mathbb{R})$ and V_0 is the closed subspace, then ϕ produces the multiresolution analysis. Therefore the scaling function is used to approximate the signal up to a particular level of detail.

A family of scaling function can be constructed via shifts and power of two stretches given by the mother scaling function.

$$\phi_{j,k} = 2^{-j/2} \phi(2^{-j}t - k) \quad j, k \in \mathbb{Z} \quad (7)$$

As with scaling function, wavelet is defined as

$$\psi_{j,k} = 2^{-j/2} \psi(2^{-j}t - k) \quad j, k \in \mathbb{Z} \quad (8)$$

In multiresolution analysis, $\phi(t)$ is a lowpass filter and $\psi(t)$ is a highpass filter. $\psi(t)$ provides localization in time and frequency.

The discrete wavelet transform was first given by Mallat. For detailed information, the reader is also referred to Mallat (Mallat, 1989).

3.3 Fast wavelet transform

The fast wavelet transform is a method to compute the discrete wavelet transform just like the fast Fourier transform which computes the discrete Fourier transform.

In the discrete Fourier transform, the transform can be made fast because the transformation is represented as a product of sparse elementary matrices. However, the transformation matrix is orthogonal and its inverse is equal to its transpose. This matrix is made orthogonal

by choosing the unit vectors as basis in time domain and the exponential terms as basis in the frequency domain. If the fast wavelet transform is generated from the discrete wavelet transform, Similar observation holds. The detailed information can be found in Mallat's book (Mallat, 1998).

4. Wavelet packet transform

The wavelet packet transform is based on the traditional discrete wavelet transform in terms of both approach and analysis of detail coefficients. With this method, better resolution can be obtained for both time and frequency depending on the contents of the data. It was first proposed by Coifman, Meyer and Wickerhauser (Coifman et al., 1992).

The wavelet transform is signal-dependent since the decomposition of the signal is performed by taking the best set of functions from the basis. The basis is determined by selecting the second filter bank tree structure, from which conversion coefficients are obtained (Coifman&Wickerhauser, 1990, 1992). Therefore, application of the separation process is simple in terms of ease of calculation.

The wavelet packet analysis for a time series can be summarized as follows (Mallat, 1989).

A space V_j of a multiresolution analysis in $L^2(\mathbb{R})$ is analyzed in a lower resolution space W_{j+1} and a detail space W_{j+1} of is added. Dividing the orthogonal basis $\{\phi_j(t-2^{jn})\}_{n \in \mathbb{Z}}$ to the new orthogonal basis constitutes $\{\phi_{j+1}(t-2^{j+1n})\}_{n \in \mathbb{Z}}$ of V_j and $\{\psi_{j+1}(t-2^{j+1n})\}_{n \in \mathbb{Z}}$ of W_{j+1}

The decompositions of ϕ_{j+1} and ψ_{j+1} are denoted by a pair of conjugate mirror filters, $h[n]$ and $g[n]$. The relation between $h[n]$ and $g[n]$ is given as (Herley&Vetterly, 1994)

Any node of the binary tree is labeled by (j, k) , where $j-L \geq 0$, j is the depth of the node on the tree, and k is the number of nodes. A space W_j^k allowing an orthonormal basis $\{\psi_j^k(t-2^{jn})\}_{n \in \mathbb{Z}}$ is associated to each node (j, k) when going down the tree. At the root, it has $W_L^0 = V_L$ and $\psi_L^0 = \phi_L$. The wavelet packet orthogonal bases at the nodes are defined by

$$\psi_{j+1}^{2k}(t) = \sum_{n=-\infty}^{+\infty} h[n] \psi_j^k(t-2^{jn}) \quad (9)$$

$$\psi_{j+1}^{2k+1}(t) = \sum_{n=-\infty}^{+\infty} g[n] \psi_j^k(t-2^{jn}) \quad (10)$$

According to this, the filter functions are derived as

$$h[n] = \langle \psi_{j+1}^{2k+1}(t), \psi_j^k(t-2^{jn}) \rangle \quad (11)$$

and

$$g[n] = \langle \psi_{j+1}^{2k+1}(t), \psi_j^k(t-2^{jn}) \rangle \quad (12)$$

where $\{\psi_j^k(t-2^{jn})\}_{n \in \mathbb{Z}}$ is orthonormal.

Because of the filtering operation in the wavelet spaces, the phase space can participate in various ways. Any choice of decompositions gives the wavelet packet decomposition. This procedure is given in Figure 6 and phase plane is shown in Figure 7.

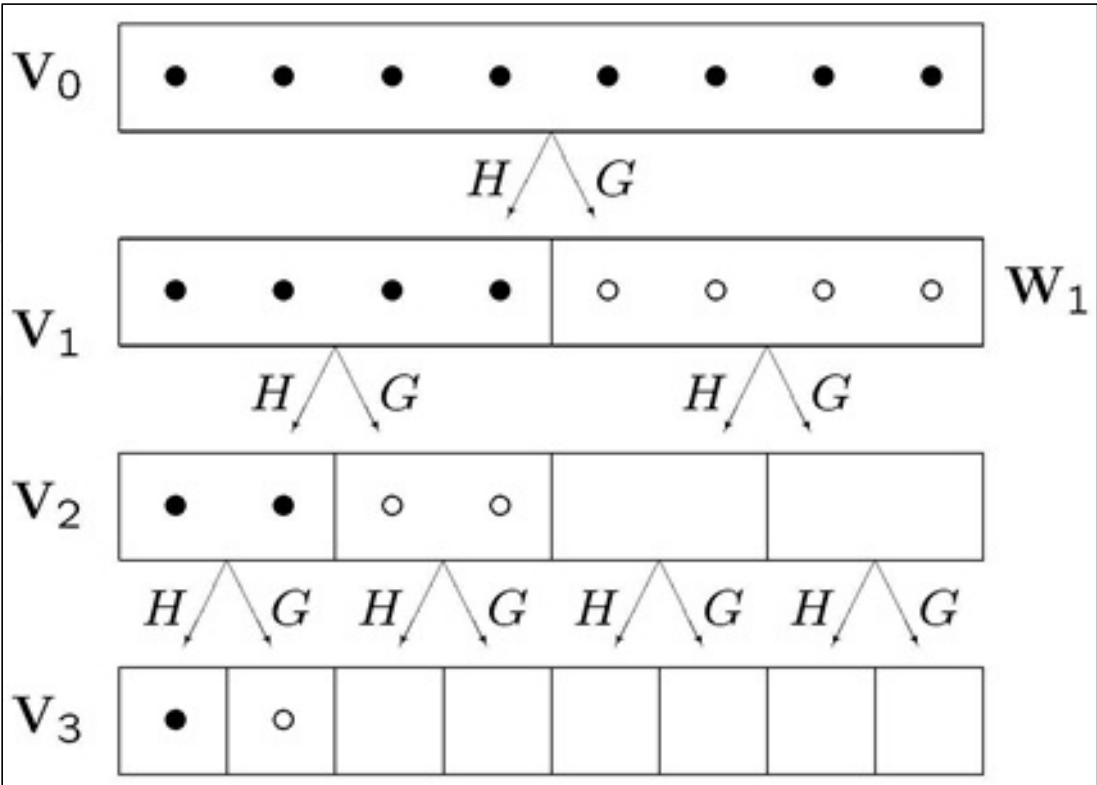


Fig. 6. General structure of the wavelet packet transform

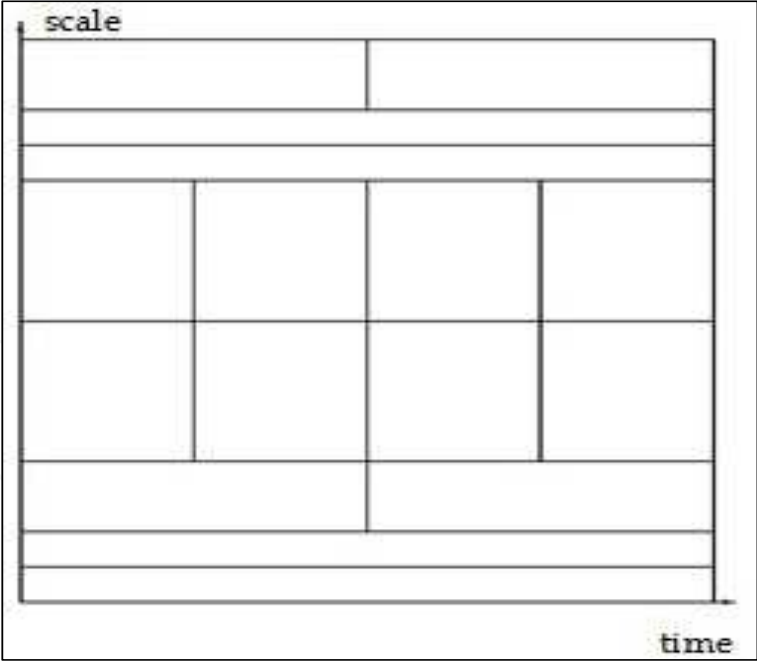


Fig. 7. The phase plane of the wavelet packet transform

From the first rule of the multiresolution analysis, the recursive splitting determines a binary tree of the wavelet packet spaces, which are defined as

$$W_{j+1}^{2k} \oplus W_{j+1}^{2k+1} = W_j^k \tag{13}$$

Consequently, coefficients of decomposition and reconstruction are

$$D_{j+1}^{2k}[t] = d_j^k * \tilde{h}[2t] \quad (14)$$

$$d_{j+1}^{2k+1}[t] = d_j^k * \tilde{g}[2t] \quad (15)$$

$$d_j^k[t] = \tilde{d}_{j+1}^{2k} * h[2t] + \tilde{d}_{j+1}^{2k+1} * g[2t] \quad (16)$$

To sub sampling of the convolution of d_j^k with \tilde{h} and \tilde{g} , the coefficients must be obtained. By the iteration of these equations, all the branches of the tree are computed by the wavelet packet coefficients. This is given in Figure 8 and Figure 9.

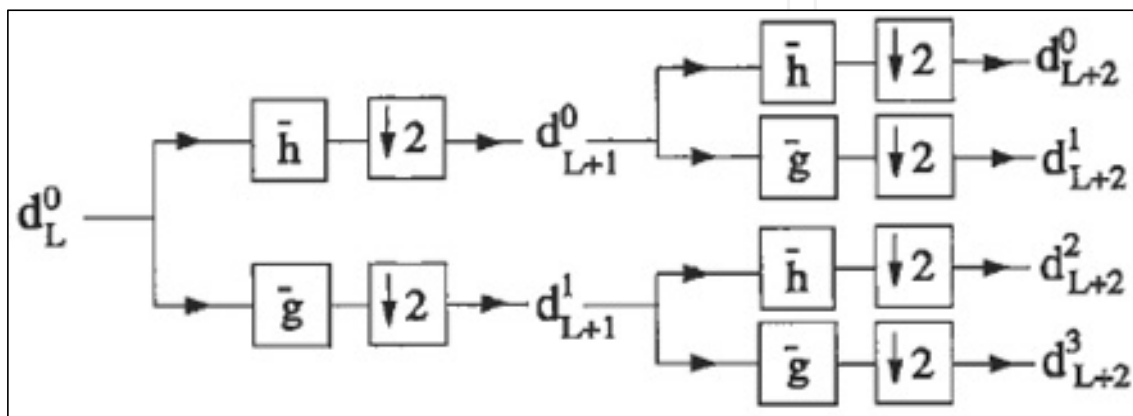


Fig. 8. Two level wavelet packet decomposition with down sampling

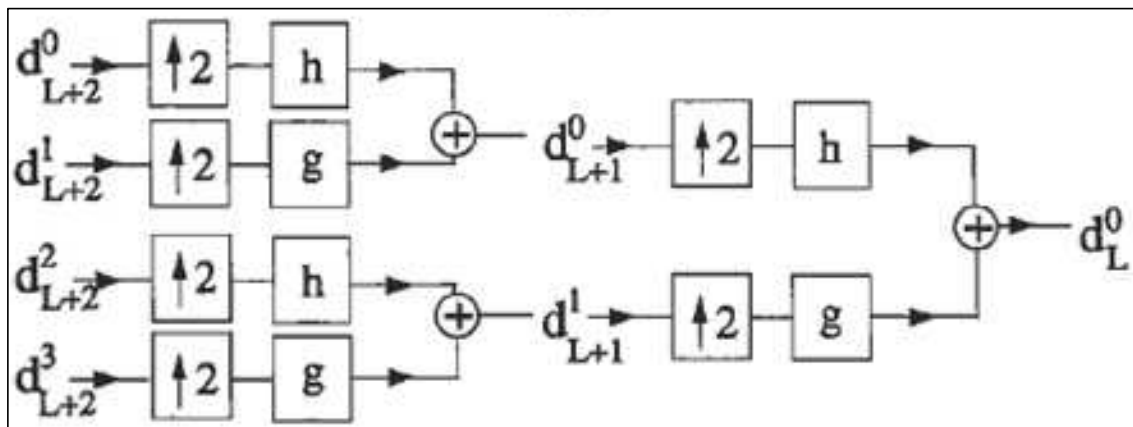


Fig. 9. Two level wavelet packet reconstruction with up sampling

Best tree function is a one- or two-dimensional wavelet packet analysis function that computes the optimal sub tree of an initial tree with respect to an entropy type criterion (Coifman&Wickerhauser, 1992). The resulting tree may be much smaller than the initial one. Following the organization of the wavelet packets library, it is natural to count the decompositions issued from a given orthogonal wavelet. A signal of length $N = 2L$ can be expanded in α different ways, where α is the number of binary sub trees of a complete binary tree of depth L , where $\alpha \geq 2^{N/2}$. This number may be very large, and since explicit enumeration is generally intractable, it is interesting to find an optimal decomposition with

respect to a convenient criterion, computable by an efficient algorithm (Donoho, 1995; Guo et al., 2000; Johnstone, 1997).

The difference between the discrete wavelet transform and the wavelet packet transform is in the decomposition of detail space. The wavelet packet transform decomposes not only the approximation space but also the detail space. This means that it can separate frequency band uniformly. Figure 10 and Figure 11 show 2-level analysis and synthesis part of the discrete wavelet transform for comparison.

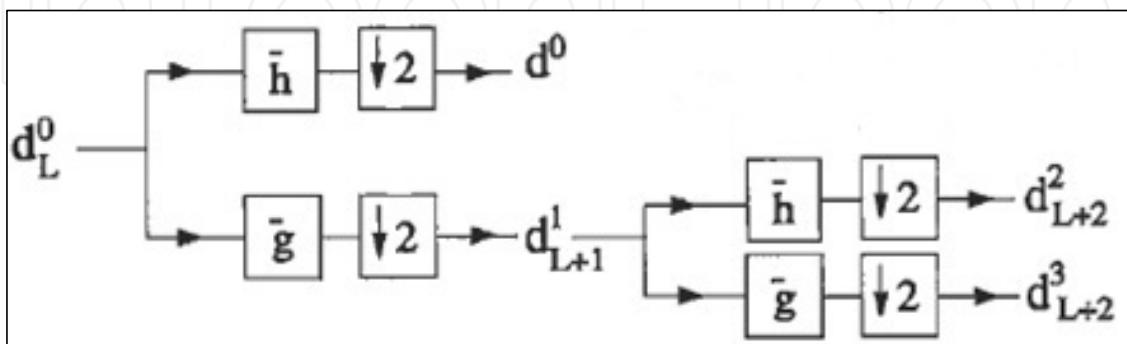


Fig. 10. Two level analysis part of the discrete wavelet transform

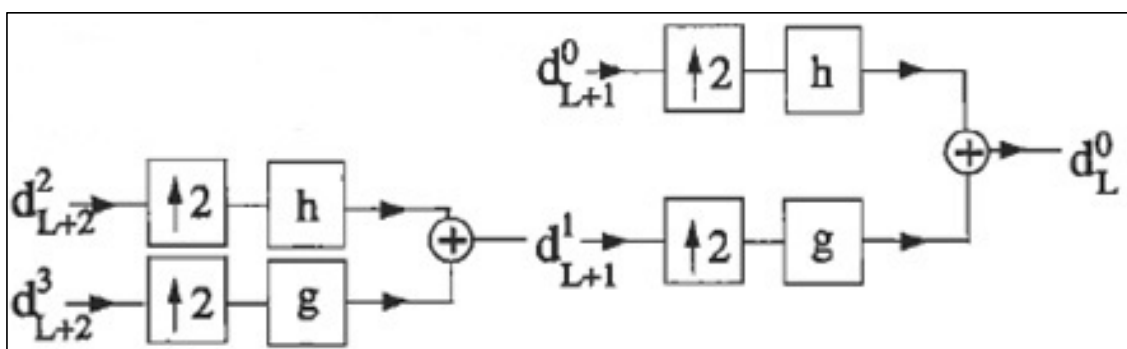


Fig. 11. Two level synthesis part of the discrete wavelet transform

5. Wavelet based speech strategy

The proposed speech processing strategy is based on wavelet packet transform. The strategy consists of five basic parts. Since the basis for the selection of electrode is frequency selection, it is improved by the use of the wavelet packet transform. The main wavelet function is experimentally selected as Daubechies 10. It is analyzed till level 8. Hanning window is used, to prevent short-term changes of the signal in windowing. A block diagram showing the proposed strategy is given in Figure 12. The channel outputs of the matching function is determined by finding the best tree in the matching block. The matching function shows the relationship between electrodes and output nodes of the wavelet packet transform. In this study the number of electrodes used to stimulate the cochlea is 22. The frequency position function is used to calculate channel frequency bands for cochlea. The selection of electrodes is the same as that of N-of-M model. 6 electrodes are selected for 22 channel electrodes. Only these 6 electrodes, with the highest amplitude, are analyzed using the wavelet packet transform.

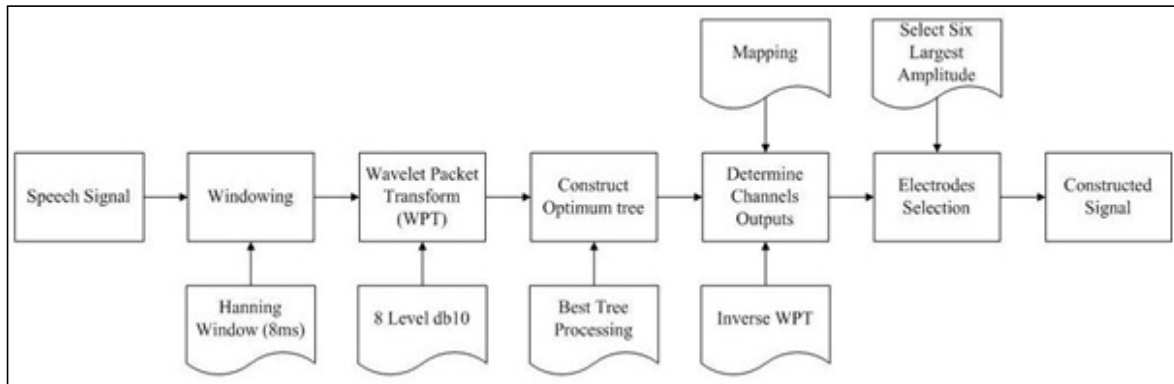


Fig. 12. Block diagram of the new speech processing method

In the new speech processing method, the best tree is determined from a block, which is structured to eliminate noise and unwanted elements of the sound signal. Thus the new output electrodes are determined with less error than those of N-of-M strategy. Moreover better transmittance is obtained by minimizing the interference between neighboring channels.

The channel outputs of the matching function is determined by finding the best tree in the matching block. The used matching function is described as

$$E_k = M \cdot C_{i,j} \quad (17)$$

Here, E_k shows the electrode output, and M shows the matching function. $C_{i,j}$ is the wavelet coefficients. The matching function shows the relationship between electrodes and output nodes of the wavelet packet transform.

In this study, 22 electrodes are used for stimulation and the frequency position function of

$$f = A \cdot 10^{ax} - k \quad (18)$$

is used to calculate channel frequency bands for cochlea (Greenwood, 1990). Here; f shows the frequency in Hz unit, while x describes the length ratio of base from 0 to 1. 'A' and 'a' are constant and their values are 156,4 and 2,1 respectively. The selection of electrodes is the same as that of N-M model. In this study, 6 electrodes are selected from 22 channel electrodes and only these 6 electrodes, which have the highest amplitude, are analyzed using the wavelet packet transform.

The proposed new strategy is essentially based on N-M strategy. The input waveform is given in Figure 13. In addition, the output waveforms for the proposed strategy and the N-M strategy are given in Figure 14 and Figure 15, showing that the input and output waveforms are very similar.

As shown from the graphs, traditional N-M strategy removes some high frequency components that are between 25 ms and 75 ms at wideband spectrogram, high frequency components are very important for intelligibility and consonant recognition such as 's', 'ş', 'f', and such. New method keeps high frequency components using the wavelet packet transform because the wavelet packet transform analyses high frequency components as well as low frequency components. Another effect is mother wavelet selection; and experimentally Daubechies 10 is found to be more effective in high frequency analysis.

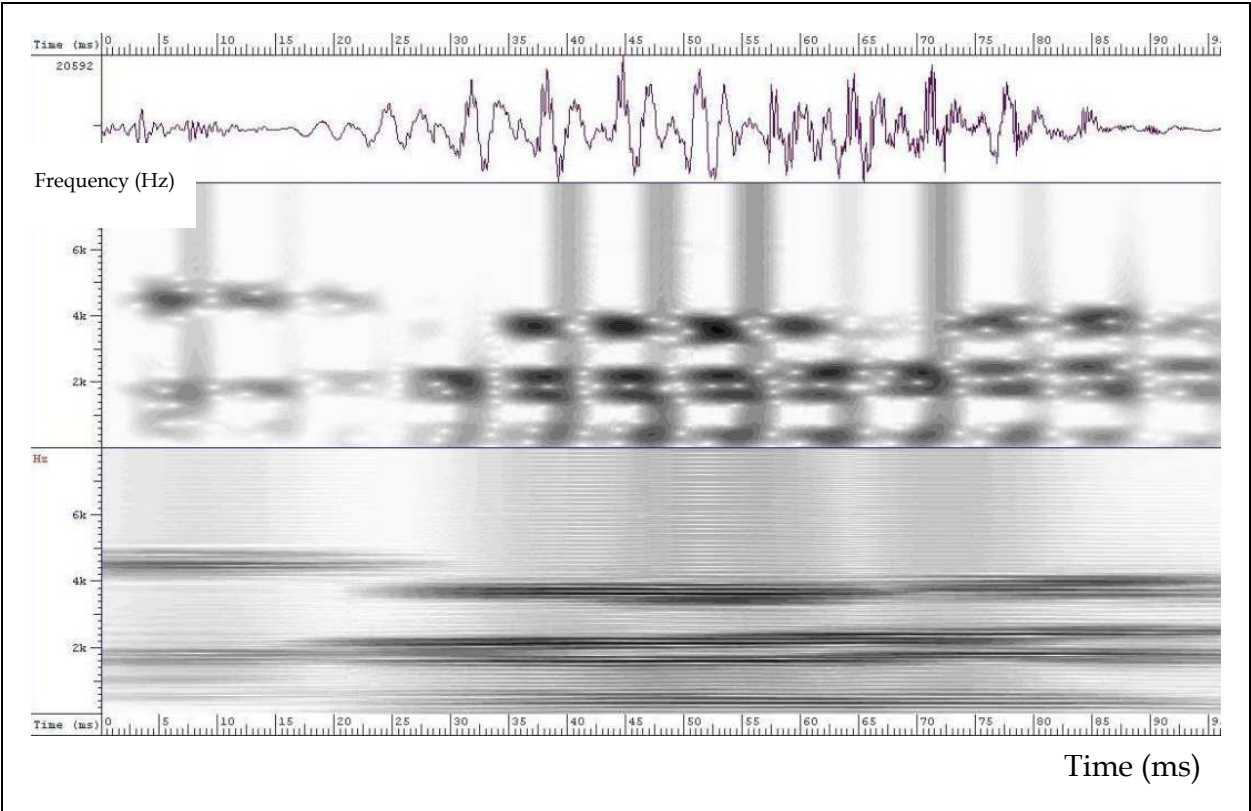


Fig. 13. Wide-band and narrow-band spectrum of the input signal

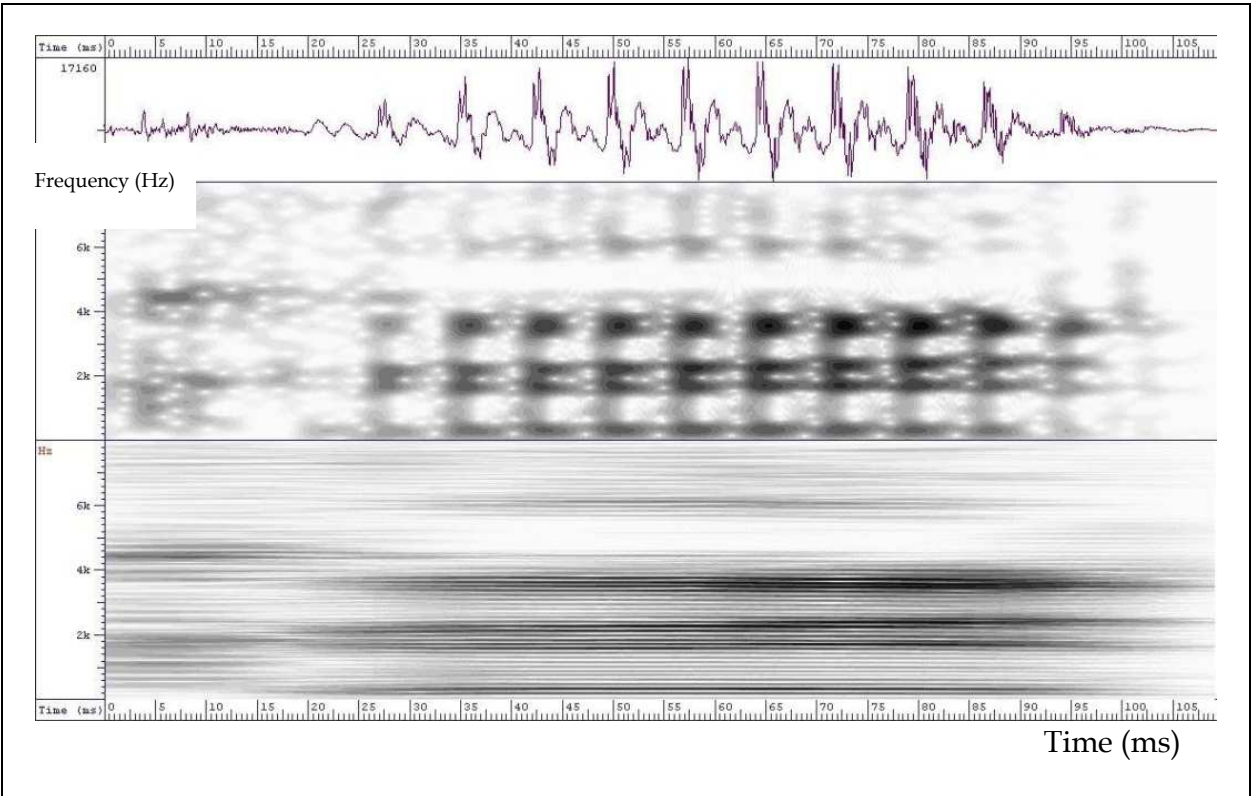


Fig. 14. Wide-band and narrow-band spectrum of the output signal for the proposed speech processing method

‘Determine optimum tree’ block eliminates noise and unnecessary components in speech signal. Therefore, it obtains better result than N-M strategy for electrode selection. New strategy output channels are more accurate than N-M strategy and conduces to reduce interaction between neighbour channels.

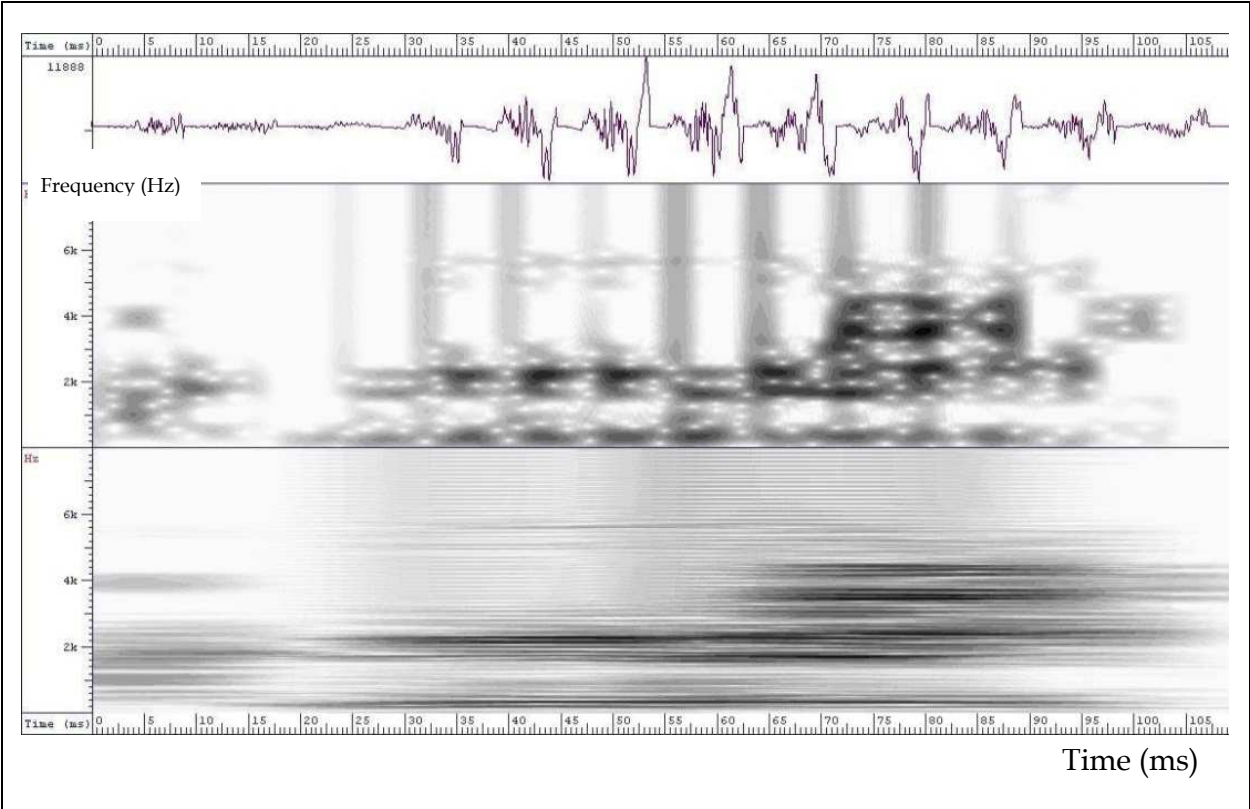


Fig. 15. Wide-band and narrow-band spectrum of the output signal for the N-M strategy

6. Experimental study

The experimental study was carried out on 20 healthy subjects between the ages of 23 and 30. The main language of all the subjects is Turkish and they have obtained air conduction thresholds better than 20 dB at octave frequencies ranging from 250 to 6000 Hz bilaterally. Since the subjects would listen to the words in both algorithms and the possibility exists that the subjects could memorize the given words and their order, separate word lists were arranged for the new strategy, and the N-of-M strategy. The arrangements were attained consultation with a Turkish Language specialist and by referring to The Turkish Language Dictionary by Turkish Language Association. All words were balanced in terms of speech knowledge and degree of difficulty at both of the lists. The usage frequencies of vowels and consonants in the lists were determined according to Turkish grammar. In addition, the lists were recorded in the silent rooms of Ege University, the Department of Otology, and the doctors of the Department approved the reader’s sound levels.

The test subjects listened to the lists from a microphone that was directly connected to their heads. While listening, they were asked to write the words they heard and leave a space for the words they could not comprehend within a table format in the listening order. The intelligibility percentage is calculated as follows.

$$\text{Intelligibility} = \frac{\text{Number of correct words}}{\text{Number of all the words on the list}} \times 100$$

(19)

The results of intelligibility test are given in Table 1 for both the N-of-M Strategy and the Proposed Speech Processing Strategy. The intelligibility percent of the proposed wavelet packet transform based strategy is higher than that of N-of-M strategy. The average percentages of intelligibility according to gender are also given in Figure 16.

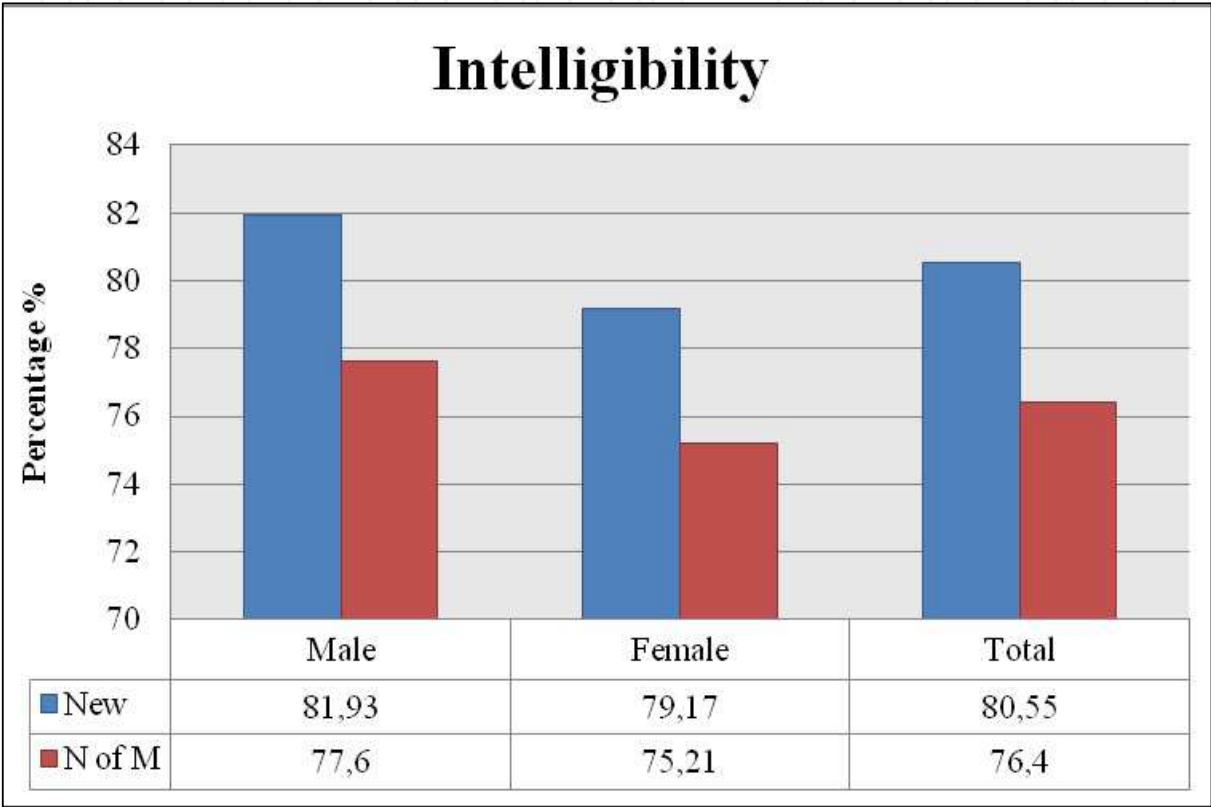


Fig. 16. Graphical representation of the intelligibility test results

	Gender	Age interval	Number of test subjects	Intelligibility (%)
New/Proposed	Male	23-29	8	81,93
	Female	23-33	12	79,17
	Total	23-33	20	80,55
N-M	Male	23-29	8	77,60
	Female	23-33	12	75,21
	Total	23-33	20	76,40

Table 1. Intelligibility test results for both the N-M strategy and the proposed strategy

To test the noise resistance of the proposed strategy, SNR improvement test was applied to the system. This test was performed by adding 5 dB of pink noise, F-16 noise, Volvo noise and factory noise respectively to the recorded data. All the added noise is real noise recorded as indicated by their names. Later, both the proposed method and the N-of-M

method were separately applied to this data. The results were evaluated by looking at the SNR given in Figure 17. Consequently, the proposed method gives better results for each type of added noise.

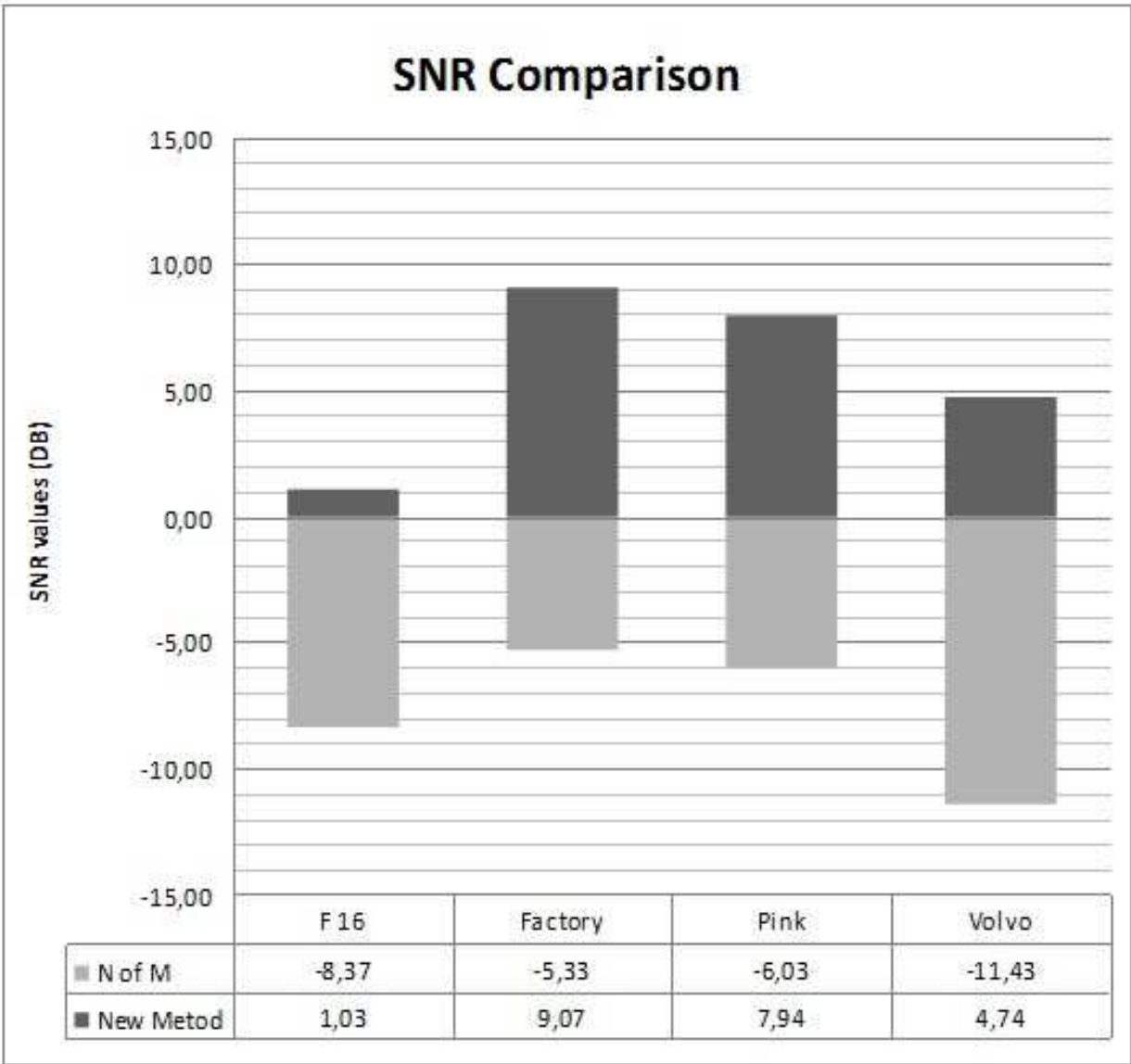


Fig. 17. SNR for different types of noise

7. Conclusion

In this study, a new speech processing strategy, based on the wavelet packet transform, is proposed for Cochlear implant applications. The foundation of the system lies in, first, obtaining the highest-energy coefficients, and then stimulating the linked electrodes and therefore improving the deaf patient’s hearing ability.

The core of the system is based on the wavelet packet transform and it also uses the energy of the wavelet coefficients. By the application of various tests, the effect of intelligibility and noise resistance for the suggested speech processing method was investigated. Then, a new

electrode selection algorithm, which depends on wavelet entropy distribution, was presented. The proposed electrode selection increases the noise performance and intelligibility. Additionally, the performance of the proposed method is better than the traditional and recently published methods.

In this study, the system was tested in terms of intelligibility; besides, SNR results were compared with those of the N-of-M Strategy. As a result, the proposed method was observed to increase performance in terms of intelligibility.

In the part of the wavelet packet transform, the determination of the optimum tree using the best tree functions is a significant part of this study because this part eliminates noise and unnecessary components from the speech signal. It also helps to improve intelligibility of speech in noisy environments.

During the live experiments session, people with normal hearing are used and also all results are based on only normal-hearing people. A further study on patients who use cochlear implant would give more accurate results for intelligibility.

For future works, hybrid mother wavelet in wavelet decomposition process is recommended. In this hybrid model, Daubechies family can be used for low-pass filter decomposition and Symlet family for high-pass filter decomposition. Moreover, the mother wavelet for deciding the choice of wavelet family can be selected according to speech signal characteristics in real time. This might give better results for speech intelligibility. Additionally, the bionic wavelet instead of the wavelet packet transform can be suggested in entire speech processing (Yuan, 2003).

8. Acknowledgment

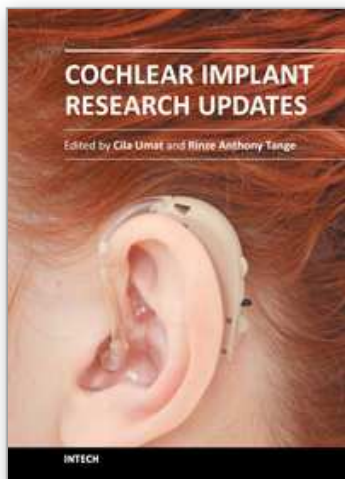
We would like to specially thank the Department of Otology in Ege University and Yahya Ozturk for their assistance in this work.

9. References

- Cheikhrouhou I., Atitallah R. B., Ouni K., Hamida A. B., Mamoudi N., Ellouze N. (2004). Speech analysis using wavelet transforms dedicated to cochlear prosthesis stimulation strategy, *Int. Sym. On Control, Communication and Signal Processing*, pp. 639-642
- Christiansen J. B., Leigh I. W. (2002). *Cochlear implants in children: Ethics and choices*, Washington, Gallaudet University Press
- Coifman R. R., Meyer Y., Wickerhauser M. V. (1992). *Wavelet Analysis and Signal Processing*
- Coifman R. R., Wickerhauser M. V. (1990). *Best adapted wave packet bases, Numerical Algorithm*, Research Group, Dept. of Mathematics, Yale University, NewHaven, Connecticut
- Coifman R. R., Wickerhauser M. V. (1992). Entropy based algorithms for best basis selection, *IEEE Trans. on Information Theory*, vol. 38, part 2
- Daubechies I. (1992). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, ISBN 0-89871-274-2
- Debnath L. (2002). *Wavelet Transforms and Their Applications*, Birhauser

- Deller J. R., Proakis J. G., Hansen J. H. L. (1994). *Discrete-time processing of speech signals*, New York, Macmillan Publishing Company
- Derbel A., Ghorbel M., Samet M., Ben Hamida A. (2009). Implementation of strategy based on auditory model based wavelet transform speech processing on DSP dedicated to cochlear prosthesis, *Int. Sym. On Comp. Intelligence and Intelligent Informatics*, pp. 143-148
- Dorman M., Hannley M., Dankowski K., Smith L., McCandless G. (1989). Word recognition by 50 patients fitted the Symbion multichannel cochlear implant, *Ear and Hearing*, vol. 10, no. 1
- Dorman M., Loizou P., Rainey D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs, *Journal of the Acoustical Society of America*, vol. 102, pp. 2403-2411
- Donoho L. (1995). Denosing by soft thresholding, *IEEE Trans. on Information Theory*, vol. 41, no. 3, pp. 613-627
- Eddington D. (1980). Speech discrimination in deaf subjects with cochlear implants, *J. Acoust. Soc. Am.*, vol. 68, no. 3, pp. 885-891, 1980
- Greenwood D. D. (1990). A cochlear frequency-position function for several species-29 years later, *J. Acoust. Soc. Am.*, vol. 87, no. 6, pp. 2593-2605
- Gopalakrishna V., Kehtarnavaz N., Loizou P. C. (2010). A recursive wavelet-based strategy for real time cochlear implant speech processing on PDA platforms, *IEEE Trans. On Biomedical Eng.*, vol. 57, no. 8, pp. 2053-2063
- Guo D., Zhu W. H., Gao Z. M., Zhang J. Q. (2000). A study of wavelet thresholding denoising, *IEEE International conference on signal processing*, vol.1, pp. 329-332
- Herley C., Vetterly M. (1994). Orthonormal time varying filterbanks and wavelet packets, *IEEE Trans. on Signal Processing*, vol. 41, no. 10
- Holschneider M. (1989). Wavelet, time frequency methods and phase space, *First International conference on wavelet*, Springer-Verlag
- House W., Urban J. (1973). Long term results of electrode implantation and electronic stimulation of the cochlea in man, *Annals of Otology, Rhinology and Laryngology*, vol. 82, pp. 504-517
- House W., Berliner K. (1982). Cochlear implants: Progress and perspectives, *Annals of Otology, Rhinology and Laryngology*, pp. 1-124
- Johnstone I. M. (1997). Wavelet threshold estimators for data with correlated noise, *J. Royal Statistical Society*, vol. 59, no. 2, pp. 319-351
- Loeb G. (1990). Cochlear prosthetics, *Annual Review in Neuroscience*, vol. 13, pp. 357-371
- Loizou P. (1998) Mimicking the human ear, *IEEE Signal Processing Magazine*, vol. 15, no. 5, pp. 101-130, 1998
- Mallat S. (1998). *A wavelet tour of signal processing*, New York, Academic Press
- Mallat S. (1989). A theory for multiresolution signal decomposition, *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 11, pp. 674-693
- Merry R. J. E. (2005) Wavelet Theory and Applications A literature study, Eindhoven Un. of Tech., Dept. of Mechanical Engineering, Control Systems Technology Group, Eindhoven, June 7, 02.05.2008 Available from <http://www.narcis.nl/publication/RecordID/oai:library.tue.nl:612762>
- Meyer Y. (1992). *Wavelets and Operators*, Cambridge University

- Millar J., Tong Y., Clark G. (1984). Speech processing for cochlear implant prostheses, *Journal of Speech and Hearing Research*, vol. 27, pp. 280-296
- Misiti M., Misiti Y., Oppenheim G., Poggi J. M. (2000) *Wavelets Toolbox Users Guide*, The MathWorks, Wavelet Toolbox
- Moore J. A., Teagle H. F. B. (2002). An introduction to cochlear implant technology, activation, and programming Language, Speech, and Hearing Services in Schools, pp.153-161
- Nogueira W., Giese A., Edler B., Buchner A. (2006). Wavelet Packet Filterbank for Speech Processing Strategies in Cochlear Implants, *ICASSP*
- Parkins C., Anderson S. (Eds.) (1983). Cochlear prostheses, *An international symposium New York*, New York Academy of Sciences
- Pereyra M. C., Mohlenkampy M. J. (2004). Wavelets, their friends, and what they can do for you, June 20, Department of Mathematics and Statistics, MSC03 2150, Un. of New Mexico, Albuquerque NM 87131-0001 USA, 23.08.2005, Available from <http://www.math.unm.edu/crisp>
- Polikar R. (1999). The wavelet tutorial, 06.06.2000, Available from <http://users.rowan.edu/polikar/WAVELETS/WTtutorial.html>
- Quatieri (2001). *Discrete-Time Speech Signal Processing*, Principles and Practice. PrenticeHall
- Saeed V. V. (2000). *Advanced Signal Processing and Noise Reduction*, 2nd Edition, John Wiley and Sons
- Rioul O., Vetterli M. (1991). Wavelets and signal processing, *IEEE SP Magazine*, pages 14-38, October 1991
- Sheikhzadeh H., Abutalebi H. R. (2001). An improved wavelet-based speech enhancement system, *Eurospeech*, pp. 1855-1858
- Strang G., Nguyen T. (1997). *Wavelets and Filter Banks*, Wellesley-Cambridge Press, second edition, ISBN 0-9614088-7-1
- Wilson B. S., Finley C. C., Lawson D. T., Wolford R. D., Eddington D. K., Rabinowitz W. M. (1991). Better speech recognition with cochlear implants, *Nature*, 352, pp. 236-238
- Wilson B., (R. Tyler, ed.) (1993). *Signal processing in Cochlear Implants: Audiological Foundations*, pp. 35-86, Singular Publishing Group, Inc.
- Yuan X. (2003). *Auditory Model-based Bionic Wavelet Transform*, Speech and Signal Processing Lab., Milwaukee, Wisconsin



Cochlear Implant Research Updates

Edited by Dr. Cila Umat

ISBN 978-953-51-0582-4

Hard cover, 232 pages

Publisher InTech

Published online 27, April, 2012

Published in print edition April, 2012

For many years or decades, cochlear implants have been an exciting research area covering multiple disciplines which include surgery, engineering, audiology, speech language pathology, education and psychology, among others. Through these research studies, we have started to learn or have better understanding on various aspects of cochlear implant surgery and what follows after the surgery, the implant technology and other related aspects of cochlear implantation. Some are much better than the others but nevertheless, many are yet to be learnt. This book is intended to fill up some gaps in cochlear implant research studies. The compilation of the studies cover a fairly wide range of topics including surgical issues, some basic auditory research, and work to improve the speech or sound processing strategies, some ethical issues in language development and cochlear implantation in cases with auditory neuropathy spectrum disorder. The book is meant for postgraduate students, researchers and clinicians in the field to get some updates in their respective areas.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Gulden Kokturk (2012). Wavelet Based Speech Strategy in Cochlear Implant, Cochlear Implant Research Updates, Dr. Cila Umat (Ed.), ISBN: 978-953-51-0582-4, InTech, Available from:
<http://www.intechopen.com/books/cochlear-implant-research-updates/wavelet-based-speech-strategy-in-cochlear-implant>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen